A Modified Fractionally Co-integrated VAR for Predicting Returns

Xingzhi Yao   Marwan Izzeldin

Department of Economics, Lancaster University

13 December 2015

Outline

1. Introduction and Motivation
2. Contribution
3. The Model
4. Monte Carlo Simulation
5. Empirical Examples
6. Concluding Remarks

Introduction

Consider a fractionally co-integrated VAR (FCVAR) model of *Johansen (2008)* and *Johansen and Nielsen (2012)*

$$\Delta^d X_t = \underbrace{\alpha\beta'\Delta^{d-b}L_b X_t}_{\text{long-run equilibrium}} + \underbrace{\sum_{c=1}^{k}\Gamma_c \Delta^d L_b^c X_t}_{\text{short-run dynamics}} + \varepsilon_t$$

- $X_t \sim I(d)$
- $\Delta^d = (1-L)^d$, $L_d = 1-\Delta^d$, ε_t is $i.i.d.(0,\Omega)$
- α: speed of adjustment, β' : the cointegrating vector, Γ_c: lag matrix
- $\beta' X_t \sim I(d-b)$, $1 > d \geq b > 0$

A wide range of applications: stock price and volatility forecasting, market return predictability, economic voting hypothesis...

Motivation

- Consider a FCVAR

$$\Delta^d X_t = \alpha\beta' \Delta^{d-b} L_b X_t + \sum_{c=1}^{k} \Gamma_c \Delta^d L_b^c X_t + \varepsilon_t$$

- $X_t = (x_t, y_t, z_t)'$, where $x_t, y_t \sim I(d)$, $z_t \sim I(0)$
  - fractional co-integration between $I(d)$ variables is affected by adding z_t, which gives rise to a biased estimate of β
  - the model is mis-specified when $d > b$: error correction terms $\beta' \Delta^{d-b} X_t$ are over-differenced

Contribution

We propose a modified FCVAR (M-FCVAR) to accommodate a mixture of $I(d)$ and $I(0)$ variables, which

- restricts shocks arising from the $I(0)$ to only have transitory effects on long-memory variables, an extension of *Fisher et al. (2015)* to fractional models
- allows for a commonly encountered case where $d \geq b$

Compared with FCVAR, M-FCVAR results in

- lower *MSEs* of co-fractional estimates
- higher degree of predictability of the $I(0)$ variable

We verify our claims using

- a monte carlo simulation
- an empirical application using high frequency financial data (SP500, SPY and VIX index) from 2003 to 2013

FCVAR vs M-FCVAR

	FCVAR	M-FCVAR
System	$X_t = (RV_t,\ VIX_t^2,\ r_t)'$	$Y_t = (\Delta^{d-b}RV_t,\ \Delta^{d-b}VIX_t^2,\ r_t)'$
Set-up	$\Delta^d X_t = \alpha\beta'\Delta^{d-b}L_b X_t + \sum_{c=1}^{k}\Gamma_c \Delta^d L_b^c X_t + \varepsilon_t$	$\Delta^b Y_t = \alpha\beta' L_b Y_t + \sum_{c=1}^{k}\Gamma_c \Delta^b L_b^c Y_t + \varepsilon_t$
β'	$\begin{pmatrix} 1 & \beta_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & \beta_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
Error correction term	$\beta'\Delta^{d-b}X_t$	$\beta' Y_t$
α	$\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{pmatrix}$	$\begin{pmatrix} 0 & 0 \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{pmatrix}$
C	$\begin{pmatrix} g & \boxed{h \quad i} \\ j & \boxed{k \quad l} \\ 0 & 0 \quad 0 \end{pmatrix}$	$\begin{pmatrix} a & \boxed{0 \quad 0} \\ d & \boxed{0 \quad 0} \\ 0 & 0 \quad 0 \end{pmatrix}$

- $RV_t,\ VIX_t^2 : CI(d,b),\ r_t \sim I(0)$
- C denotes the long-run responses of each variable to shocks in ε_t and $\beta' C = 0_{2\times 3}$
- $\varepsilon_t = (\varepsilon_{1t}^P,\ \varepsilon_{2t}^T,\ \varepsilon_{3t}^T)'$: one permanent shock and two transitory shocks
- ideally, $C = \begin{pmatrix} a & \boxed{\begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix}} \\ d & \\ 0 & 0 \quad 0 \end{pmatrix}$
- To illustrate the restrictions on α, we introduce the Impulse-Response Functions (IRFs).

M-FCVAR: IRFs

Consider IRFs to investigate the model-implied dynamic dependencies

- $\Delta^d = (1-L)^d = \sum_{i=0}^{\infty} \theta_i(d) L^i$ with $\theta_i(d) = (-1)^i \binom{d}{i}$;
  $L_d = 1 - \Delta^d = -\sum_{i=1}^{\infty} \theta_i(d) L^i$
- The model can be written as

$$
\begin{aligned}
Y_t &= \sum_{i=1}^{\infty} (-I\theta_i(b) - \alpha\beta'\theta_i(b) + \sum_{c=1}^{k} (-1)^c \Gamma_c \mathbf{K}_{c,i}) L^i Y_t + \varepsilon_t \\
&= \sum_{i=1}^{\infty} \Xi_i L^i Y_t + \varepsilon_t
\end{aligned}
$$

  where $\mathbf{K}_{c,i} = \Delta^b L_b^c = \theta_l(b) \mathbf{K}_{c-1,i-l}$ $(l = 1, 2, 3, \cdots)$
- Invert the AR polynomial Ξ_i to obtain the IRF coefficient Φ_j

$$
\Phi_0 = I,\ \Phi_j = \sum_{i=0}^{j-1} \Xi_{j-i} \Phi_i
$$
$$
Y_t = \sum_{j=0}^{\infty} \Phi_j \varepsilon_{t-j}
$$

M-FCVAR: transitory shocks from the I(0)

Fractional co-integration is unaffected if the shocks arising from the $I(0)$ are transitory

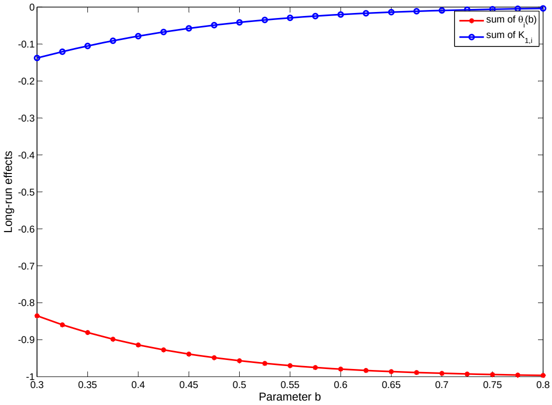
- Let $RV_t^* = \Delta^{d-b} RV_t$, which is the only variable containing permanent shocks within the model

$$
\begin{aligned}
RV_t^* &= L_b RV_t^* + \alpha_{11} L_b \xi_t + \alpha_{12} L_b r_t + (\Gamma_{11}^1 - \frac{1}{\beta_1}\Gamma_{12}^1) L_b \Delta^b RV_t^* + \frac{1}{\beta_1}\Gamma_{12}^1 L_b \Delta^b \xi_t \\
&\quad + \Gamma_{13}^1 L_b \Delta^b r_t + \varepsilon_t \\
&= -\sum_{i=1}^{\infty} \theta_i(b) L^i RV_t^* - \alpha_{11} \sum_{i=1}^{\infty} \theta_i(b) L^i \xi_t - \alpha_{12} \sum_{i=1}^{\infty} \theta_i(b) L^i r_t + (\Gamma_{11}^1 - \frac{1}{\beta_1}\Gamma_{12}^1) \sum_{i=1}^{\infty} \mathbf{K}_{1,i} L^i RV_t^* \\
&\quad + \frac{1}{\beta_1}\Gamma_{12}^1 \sum_{i=1}^{\infty} \mathbf{K}_{1,i} L^i \xi_t + \Gamma_{13}^1 \sum_{i=1}^{\infty} \mathbf{K}_{1,i} L^i r_t + \varepsilon_{t1}
\end{aligned}
$$

where $\xi_t = \Delta^{d-b}(RV_t + \beta_1 VIX_t^2) \sim I(0)$ and $r_t \sim I(0)$.

- Shocks associated with ξ_t and r_t are expected to have no long-run effects on RV_t^*
  - $\alpha_{11} = \alpha_{12} = 0$
  - no restrictions on Γ_{12}^1, Γ_{13}^1

M-FCVAR: restrictions on Alpha



- Long-run effects contributed by the lag components are negligible or even zero once b exceeds 0.6.
- Sufficient to set $\alpha_{11} = \alpha_{12} = 0$ to ensure that shocks from $I(0)$ error correction terms produce no permanent effects on RV_t^*.

Predictive R-square Implied by the M-FCVAR

- $e3' \equiv (0,0,1)$ and Φ_j denotes the IRF coefficient

$$r_t = e3' \sum_{i=0}^{\infty} \Phi_i \varepsilon_{t-i}$$

- Decompose return into the expected and unexpected part

$$r_t^h = \underbrace{e3' \sum_{j=0}^{h-1} \sum_{i=j+1}^{\infty} \Phi_i \varepsilon_{t+j-i}}_{\text{expected (A)}} + \underbrace{e3' \sum_{j=0}^{h-1} \sum_{i=0}^{j} \Phi_i \varepsilon_{t+j-i}}_{\text{unexpected (B)}}$$

- Return predictability over h horizons implied by the M-FCVAR model

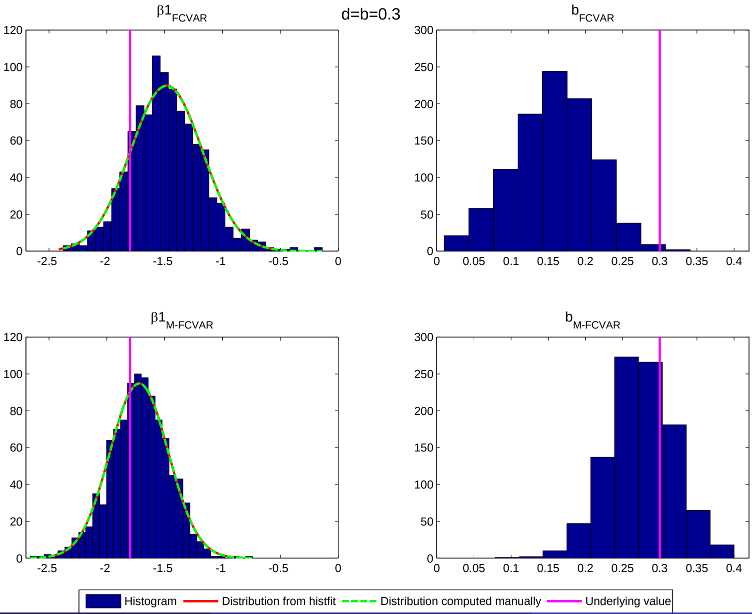
$$R_h^2 = \frac{var(A)}{var(A)+var(B)}$$

Simulation Design and Settings

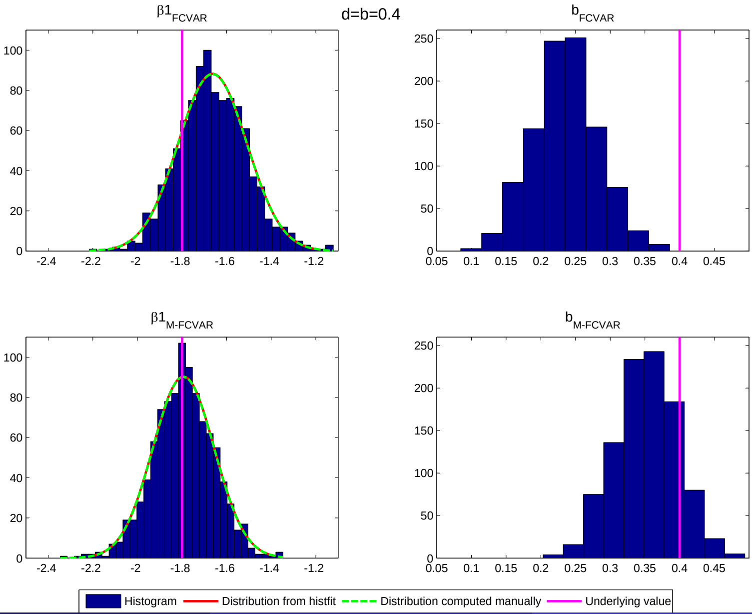
- To highlight the gains of M-FCVAR, we simulate a set of fractionally co-integrated systems containing both $I(d)$ and $I(0)$ variables
- $\varepsilon_t \sim i.i.d.N(0,1)$
- System 1: X_t
  - $(x_t,\, y_t)' = \sum_{j=0}^{\infty} \Phi_j \varepsilon_{t-j},\ \Phi_i \sim (\alpha, \beta, \Gamma, b)$
  - $X_t = (\Delta^{b-d} x_t,\, \Delta^{b-d} y_t)',\ X_t \sim CI(d,\, b)$
- System 2: Y_t
  - $Y_t = \sum_{j=0}^{\infty} \Phi_j^* \varepsilon_{t-j},\ \Phi_i^* \sim (\alpha^*, \beta^*, \Gamma^*, b)$
  - $z_t = e3' Y_t,\ z_t \sim I(0)$
- System 3: $Z_t = (X_t,\, z_t)'$
- Estimate Z_t by FCVAR and M-FCVAR, respectively. The distribution of estimates is obtained by 1000 replications of a moving block bootstrap.

$T = 2000$	α	β'	k	Γ
System 1: X_t	$\begin{pmatrix} -0.500 \\ 0.800 \end{pmatrix}$	$\begin{pmatrix} 1 & -1.800 \end{pmatrix}$	1	$\begin{pmatrix} 1.240 & -0.120 \\ -0.550 & 0.920 \end{pmatrix}$
	α^*	$\beta^{*\prime}$	k	Γ^*
System 2: Y_t	$\begin{pmatrix} -0.500 & 0.010 \\ 0.800 & -0.018 \\ -0.002 & -0.900 \end{pmatrix}$	$\begin{pmatrix} 1 & -1.800 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	1	$\begin{pmatrix} 1.240 & -0.120 & 0 \\ -0.550 & 0.920 & 0.008 \\ 0 & 0 & 0 \end{pmatrix}$

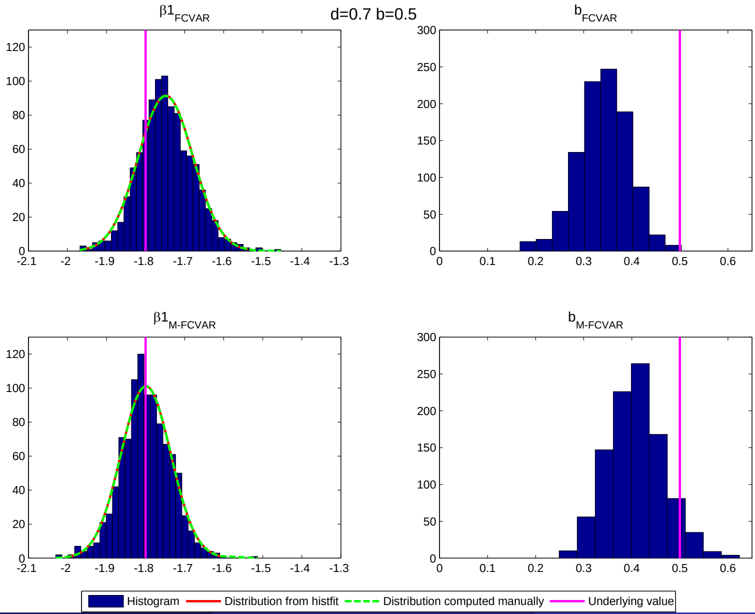
Empirical Distributions: d=b=0.3



Empirical Distributions: d=b=0.4



Empirical Distributions: d=0.7 b=0.5



Simulation Results: parameter b

		$\widehat{b}_{FCVAR}$			$\widehat{b}_{M-FCVAR}$			MSE Reduction
d	b	Bias	Std	MSE	Bias	Std	MSE	
0.3	0.3	0.143	0.054	0.023	0.026	0.044	0.003	86.957%
0.4	0.3	0.126	0.061	0.019	0.002	0.044	0.002	89.474%
0.4	0.4	0.164	0.047	0.029	0.048	0.045	0.004	86.207%
0.5	0.3	0.105	0.064	0.015	0.020	0.044	0.002	86.667%
0.5	0.4	0.151	0.049	0.025	0.044	0.043	0.004	84.000%
0.5	0.5	0.180	0.051	0.035	0.092	0.056	0.011	68.571%
0.6	0.3	0.008	0.068	0.011	0.019	0.046	0.002	81.818%
0.6	0.4	0.136	0.056	0.022	0.044	0.045	0.004	81.818%
0.6	0.5	0.170	0.053	0.032	0.092	0.053	0.011	65.625%
0.6	0.6	0.158	0.061	0.035	0.120	0.069	0.019	45.714%
0.7	0.3	0.041	0.075	0.007	0.020	0.045	0.002	71.429%
0.7	0.4	0.115	0.057	0.016	0.045	0.043	0.004	75.000%
0.7	0.5	0.159	0.054	0.028	0.091	0.054	0.011	60.714%
0.7	0.6	0.169	0.065	0.033	0.120	0.065	0.019	42.424%

Simulation Results: Beta

d	b	$\widehat{\beta}_{1}$ _FCVAR Bias	Std	MSE	$\widehat{\beta}_{1}$ _M−FCVAR Bias	Std	MSE	MSE Reduction
0.3	0.3	0.316	0.311	0.196	0.079	0.252	0.070	64.286%
0.4	0.3	0.323	0.361	0.235	0.077	0.294	0.092	60.851%
0.4	0.4	0.137	0.155	0.042	0.006	0.137	0.019	54.762%
0.5	0.3	0.279	0.306	0.207	0.078	0.294	0.093	55.072%
0.5	0.4	0.161	0.198	0.065	0.004	0.170	0.028	56.923%
0.5	0.5	0.059	0.073	0.009	0.003	0.063	0.004	55.556%
0.6	0.3	0.233	0.349	0.176	0.069	0.304	0.097	44.886%
0.6	0.4	0.145	0.204	0.062	0.004	0.160	0.025	59.677%
0.6	0.5	0.054	0.073	0.008	0.003	0.063	0.004	50.000%
0.6	0.6	0.032	0.068	0.006	0.018	0.062	0.004	33.333%
0.7	0.3	0.157	0.353	0.149	0.080	0.294	0.093	37.584%
0.7	0.4	0.118	0.204	0.056	0.007	0.171	0.029	48.214%
0.7	0.5	0.051	0.070	0.008	0.001	0.064	0.004	50.000%
0.7	0.6	0.031	0.066	0.005	0.018	0.064	0.004	20.000%

Simulation Results: predictive R-square

d	b	R^2_{FCVAR} (%) h=1	h=5	h=22	h=100	$R^2_{M-FCVAR}$ (%) h=1	h=5	h=22	h=100
0.3	0.3	0.294	0.446	0.588	0.646	0.345	0.776	1.555	3.261
0.4	0.3	0.401	0.629	0.850	0.957	0.481	1.137	2.320	4.886
0.4	0.4	0.407	0.442	0.530	0.521	0.498	1.104	2.477	5.976
0.5	0.3	0.147	0.083	0.086	0.083	0.155	0.143	0.235	0.428
0.5	0.4	0.500	0.686	0.921	0.957	0.668	1.721	3.818	8.481
0.5	0.5	0.554	0.421	0.471	0.402	0.637	1.157	2.562	5.682
0.6	0.3	0.241	0.179	0.139	0.089	0.238	0.163	0.148	0.164
0.6	0.4	0.246	0.112	0.118	0.107	0.270	0.275	0.530	1.117
0.6	0.5	0.610	0.748	0.970	0.860	0.849	2.157	4.667	9.148
0.6	0.6	0.721	0.390	0.404	0.279	0.758	1.004	1.989	3.282
0.7	0.3	0.335	0.646	0.680	0.603	0.434	1.293	2.112	3.348
0.7	0.4	0.357	0.189	0.148	0.102	0.357	0.213	0.283	0.520
0.7	0.5	0.366	0.123	0.122	0.094	0.398	0.336	0.642	1.242
0.7	0.6	0.728	0.793	0.952	0.664	0.996	2.244	4.315	6.202

Note: shocks, taken from the residuals, have been orthogonalised.

Empirical Application

- Daily CBOE VIX volatility index
- 5 min observations of aggregate S&P 500 composite index, SPDR S&P 500 ETF TRUST (SPY) index
- Sample period: 2003−2013 (2623 obs)
- 5 min intraday return $r_{t,j} = 100(\log p_{t,j} - \log p_{t,j-1})$, daily return $r_t = \sum_{j=1}^{M} r_{t,j}$
- Daily realised variance

$$rv_t = \sum_{j=1}^{M} r_{t,j}^2$$

- One-month forward realised return variation

$$RV_t = \log\left(\sum_{i=1}^{22} rv_{t+i}\right)$$

- Risk-neutral return variation

$$VIX_t^2 = \log\left(\frac{30}{365}\left(VIX_t^{CBOE}\right)^2\right)$$

FCVAR vs M-FCVAR

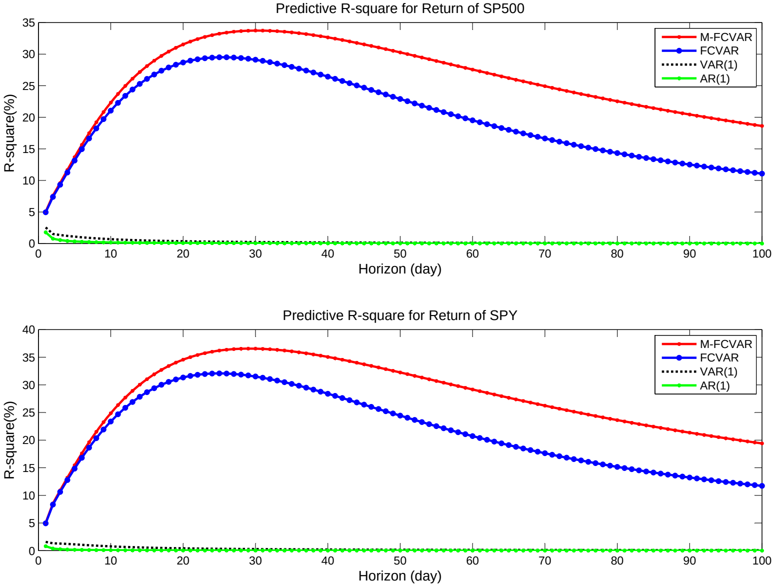
M-FCVAR: $Y_t = (\Delta^{d-b}RV_t, \Delta^{d-b}VIX_t^2, r_t)$

	d	b	k	α	β	AIC	BIC
SP500	0.694 (0.000)	0.646 (0.018)	1	$\begin{pmatrix} \underset{(0.000)}{0.000} & \underset{(0.000)}{0.000} \\ \underset{(0.013)}{0.107} & \underset{(0.038)}{0.023} \\ \underset{(0.071)}{-0.391} & \underset{(0.271)}{-1.366} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ -\underset{(0.001)}{0.974} & 0 \\ 0 & 1 \end{pmatrix}$	-6513	-6390
SPY	0.681 (0.000)	0.646 (0.015)	1	$\begin{pmatrix} \underset{(0.000)}{0.000} & \underset{(0.000)}{0.000} \\ \underset{(0.012)}{0.118} & \underset{(0.061)}{0.016} \\ \underset{(0.086)}{-0.550} & \underset{(0.494)}{-1.320} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ -\underset{(0.001)}{0.936} & 0 \\ 0 & 1 \end{pmatrix}$	-6075	-5952

FCVAR: $X_t = (RV_t, VIX_t^2, r_t)$

	d	b	k	α	β	AIC	BIC
SP500	0.694 (0.000)	0.641 (0.020)	1	$\begin{pmatrix} \underset{(0.005)}{-0.009} & \underset{(0.024)}{0.006} \\ \underset{(0.015)}{0.114} & \underset{(0.066)}{0.020} \\ \underset{(0.073)}{-0.378} & \underset{(0.675)}{-1.503} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ \underset{(0.001)}{-0.953} & 0 \\ 0 & 1 \end{pmatrix}$	-6513	-6378
SPY	0.681 (0.000)	0.635 (0.021)	1	$\begin{pmatrix} \underset{(0.005)}{-0.012} & \underset{(0.034)}{0.005} \\ \underset{(0.017)}{0.128} & \underset{(0.094)}{0.013} \\ \underset{(0.096)}{-0.569} & \underset{(1.119)}{-1.434} \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ \underset{(0.000)}{-0.916} & 0 \\ 0 & 1 \end{pmatrix}$	-6078	-5943

Predictive R-square: different models



Conclusion

- We provide modifications for the traditional FCVAR model of *Johansen (2008)* and *Johansen and Nielsen (2012)*,
- Our modified FCVAR (M-FCVAR) is better suited for the estimation of fractionally co-integrated systems containing a mixture of $I(d)$ and $I(0)$ variables.
- Specifically, the M-FCVAR
  - restricts shocks emanating from the $I(0)$ variable to only have transitory effects on the $I(d)$ variables.
  - accounts for long memory in the co-integration residuals by applying a partial differencing procedure.
  - provides co-fractional estimates with lower *MSEs*.
  - yields higher degree of predictability of the $I(0)$ variable.
- We demonstrate the gains from adopting the M-FCVAR via Monte Carlo simulations and an empirical application using high frequency data.